A Multi-Tier Cluster Based Tracker Approach for Battlefield Acoustic Systems

Sept 30, 2000

Brian T. Mays
Joseph Price
U.S. Army Research Laboratory
Adelphi, MD 20783-1197

ABSTRACT

The Acoustic Signal Processing Branch of the U.S. Army Research Laboratory (ARL) has ongoing research into Battlefield target tracking. The classic approach is to combine information from multiple line-of-bearing (LOB) sensors that are spatially diverse. The triangulations form candidate intersection points that a tracking algorithm can de-ghost and develop track histories on. For these traditional trackers the vehicles of interest must be resolved by multiple sensors simultaneously to form a valid intersection, which is attainable in sparse vehicle scenarios. In the scenario of an active Battlefield, the sensor field performance can be radically different, the sensors will be captured by their nearest targets and lose the ability to produce valid LOB intersections across the field due to each sensor hearing a different target. This capture effect will force traditional trackers to fail. This paper will develop the concept of a multi-tier tracker, which works at micro level and a macro level. At the micro level the sensor field will focus on producing an accurate estimate of vehicle count and rough estimate of cluster geometry. The cluster estimate produced does not require the simultaneous vehicle resolution by the sensors. This cluster estimate can then be tracked at the macro level via a traditional tracker. The cluster estimation and tiered tracking will provide robust theater level tracker operation with realistic sensor performance.

Introduction

Passive acoustic sensors have come a long way from being simple detection devices. Advancements in audio analysis, adaptive beamforming, prior-knowledge filtering, and terrain-based reasoning have led to a demonstrated ability to track and classify multiple vehicles traveling in widely spaced convoys. But significant additional work needs to be done to attain an Objective Force capability to track multiple, closely spaced targets on an open battlefield. The need for additional work stems from the fact that traditional acoustic sensor systems used for target location, work on the simple principle of LOB intersection from at least two sensors. When there is only one target of interest the intersection is unique and no additional problems arise as long as two sensors detect the target at all times. When this type of system is presented with multiple targets several system level degradations occur. In the case of the optimum sensor

Form SF298 Citation Data

Report Date *("DD MON YYYY")* 30092000	**Report Type** N/A	**Dates Covered (from... to)** *("DD MON YYYY")*

Title and Subtitle A Multi-Tier Cluster Based Tracker Approach for Battlefield Acoustic Systems	**Contract or Grant Number**
	Program Element Number
Authors Mays, Brian T.; Price, Joseph	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) U.S. Army Research Laboratory Adelphi, MD 20783-1197	**Performing Organization Number(s)**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Monitoring Agency Acronym**
	Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Document Classification unclassified	**Classification of SF298** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** unlimited
Number of Pages 10	

system, the sensors will still produce LOBs to all targets in the field. Under these conditions the LOB intersections will contain the true target locations as well as ghost targets. In general the ghost targets can not be immediately eliminated and must be dealt with. The next section discusses how a system can still produce useful target information under these conditions. In the case of poor sensor performance, each sensor will be captured by its closest target. Since all sensors will have a different target as their closest object, LOBs from multiple sensors only contain ghost targets and no true targets can be resolved. While typical sensor performance will fall somewhere between these two cases, the standard assumption that sensors will simultaneously resolve all targets in an active battlefield will cause system level failure.

This paper will discuss concepts for tracking battlefield movement when individual targets cannot be resolved. The concept of an organized cluster of targets becomes the focus when combining sensor measurements. The cluster is termed organized since in the battlefield scenario vehicles will not in general be moving in random directions nor have random orientation to each other. Vehicles will tend to travel in an organized fashion and be grouped together. The cluster estimate will have a target count as well as a shape estimate. Once the cluster estimate is formed it can then be tracked by a traditional tracker. While individual targets may not be resolvable, useful information such as group movement and amassment may be determined by monitoring the clusters as they travel at the theater level. This technique will allow the acoustic based systems to provide valuable information in the most challenging battlefield environments.

The assumed level of sensor performance is that the sensors can resolve multiple targets but only those which are the closest targets to themselves. This also implies that the sensors can produce vehicle counts and angular directions during pass-bye events.

The problem of ghosts

A typical acoustic sensor array is comprised of (n) microphones equally spaced about the circumference of a circle. Provided there is sufficient spectral or spatial separation of targets, each sensor array is capable of resolving a maximum of (n-1) targets. A line-of-bearing ("LOB") is then calculated for each detected target and the intersections of these LOBs with those taken simultaneously from other sensor arrays represent potential target locations.

Now consider the scenario depicted in **Figure 1** in which two sensor arrays (identified by the small concentric circles) generate LOBs to two separate targets. Without additional information it is not known whether the targets are located at the intersections represented by the solid dots or by the open circles. The additional information needed could be obtained from a priori knowledge, from advanced filtering techniques, from monitoring the intersection point dynamics, or from additional sensors.

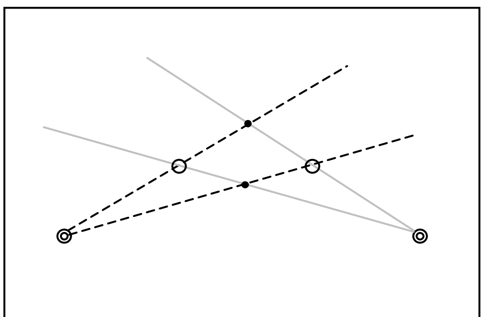

Figure 1. Which intersections are the targets?

As an example of using a priori knowledge, if one target enters the surveillance area before the other, then its established track will automatically determine the location of the second because the both targets *must* either be dots or circles, exclusively otherwise there would be fewer lines-of-bearing. Another type of a priori knowledge is geographic information--the known locations of roads and obstructions. Advanced filtering techniques include information weighting and modeling. For example, sound propagation models provide reasonable range estimations when the weather conditions and terrain are favorable; otherwise the sound path(s) become strongly and adversely affected. The monitoring of intersection point dynamics helps to discriminate targets because “ghosts” (false targets) often generate improbable velocities or accelerations. Finally, additional sensors that could be used, range in complexity from very localized “trip-wire-function” detectors, to advanced imaging devices.

In recent tests, the ability to track multiple vehicles traveling in widely spaced convoys was demonstrated using a combination of all of the above techniques. In particular, a greater number of acoustic sensors were employed to resolve targets by providing additional LOBs. This strategy had the added benefits of providing for both a larger surveillance area and for redundancy in the event of equipment failure. Development of this strategy for the open battlefield poses a significant challenge. As additional LOBs are added to confirm targets, the number of ghosts grows exponentially until it may no longer be possible to resolve ghosts from true targets at all.

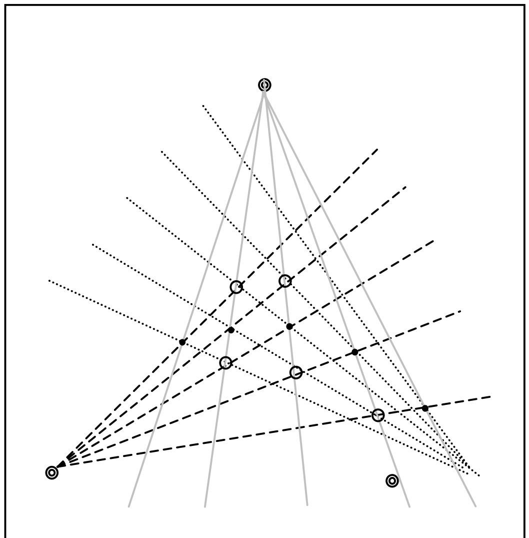

Figure 2: Triangular formation of 3 sensor arrays. 10 potential targets, 5 true targets.

An illustration of this point is given in **Figure 2**. Assuming that all targets are detected by all sensors the number of LOB intersections that will be produced is given by $\mathbf{n^2\, m\, (m - 1)\, /2}$, where **n** is the number of targets and **m** is the number of sensors. In this example, there are 3 sensors and each sensor has determined an *exact* LOB to 5 true targets located at the dots. (As mentioned above, this implies that each sensor has at least 6 microphones.) Notice that if the distance between the sensors is on the order of 500 meters, then the small circles, through which 3 LOBs cross, represent 10 meter diameters. Now, because real acoustic sensors have measurement errors, they do not determine *exact* LOBs. In practice, the accommodation made for this uncertainty is at least the same order of magnitude as the diameter of the small circles. Consequently, in the absence of additional information the small circles will be tracked as true targets. Moreover, attempts to resolve this problem by adding yet another sensor array for final confirmation becomes self defeating because the potential target density will increase along with the exponential growth in the number of LOB intersections. (In general, if a 4th sensor is added to the arrangement of Figure 2, then intersections will be formed in which 4 LOBs cross within 10 meter circles.). This example assumed that all targets where resolved by all sensors. If one assumes that some sensors do not see all the targets, then all of the LOB pairs must be treated as potential targets. This may also have the effect that some targets may not be resolved by the given LOB set if at least two sensors do not produce LOBs to such targets.

The problem of LOB formulation with typical sensor performance

So far, it has been implied that each acoustic sensor array is able to see every target. The caveat stated for the ghosting problem applies here as well: "Provided there is sufficient spectral or spatial separation of targets, each sensor array is capable of resolving a maximum of (n-1) targets." In reality, even for optimal case in which targets are widely spaced and traveling in a straight line, vehicles are apt to mask one another. (**Figure 3**[1]) Moreover, on an active battlefield where vehicles are maneuvering, the individual microphones are likely to become saturated by the audio power spectrum of the nearest targets and thereby lose the ability to produce valid LOB intersections to other targets located across the field. In the absence of information from other types of sensors (e.g. optical, seismic, or magnetic), this "capture effect" will force traditional trackers to fail.

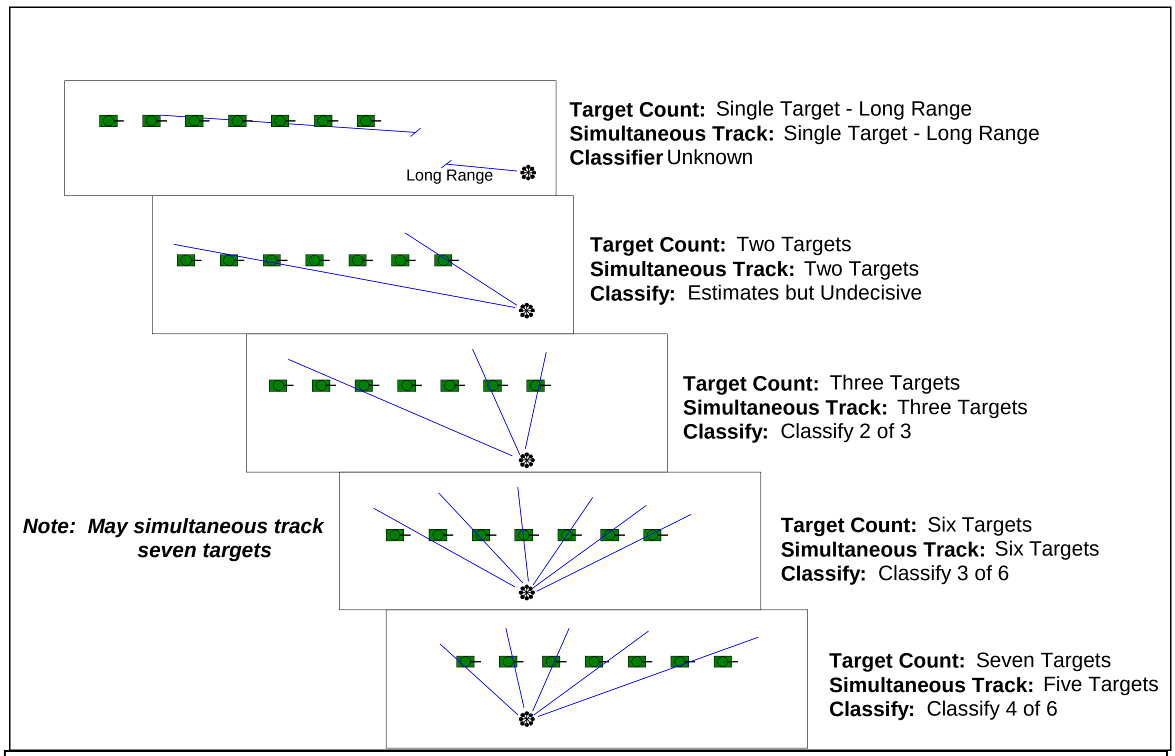

Figure 3. Target tracking scenario during vehicle approach. Expected target tracking performance is shown as multiple targets approach the sensor. (From Reference 1)

A Complementary Hierarchical Approach

As the previous sections have shown, individual target resolution may be a formidable task in the active battlefield scenario. A more robust method would be using a hierarchical approach to estimating the targets in a given area. This approach would use standard localization and de-ghosting techniques when the targets are sparse and switch to cluster based tracking when individual targets are not resolvable. Once in the cluster mode, the primary estimates are on target number, cluster shape and cluster motion throughout a theater of interest. Under the most severe conditions, only target numbers and motion may be inferred and the cluster size only bounded by the interior region of the detecting sensors.

Consider the 3 sensor array case shown in **Figure 2** again. From the formula $\mathbf{n^2\ m\ (m - 1)\ /2}$ one knows that if the number of intersections is between 1 and 3, then *at least* 1 potential target is present, even if one of the sensor arrays has become captured by a second target. Extending this argument to the 5 target scenario, (i.e. the maximum number of targets that can be resolved with 6-microphone arrays) one obtains **Table 1**. For example if we know there are 46 intersections, we can deduce that *at least* 5 targets are present, even if we don't know where they are.

Table 1: Interpretation of intersections generated by 3 acoustic sensor arrays

Intersections	*1 - 3*	*4 – 12*	*13 – 27*	*28 – 45*	*46 – 75*
Targets Required	1	2	3	4	5

Now consider the placement of two additional sensor arrays, A and B, as shown in **Figure 4**. If sensor arrays at S1, S2 and S3 are detecting targets in the interior area formed by their respective baselines (S1-S2-S3) then the gray area illustrates the region with target activity. Now this region can be established as an initial cluster. As the targets move through the gray area each sensor will be able to determine the sign of the angular rate of the cluster (CW or CCW). By examining the rotation directions as the LOBs approach a boundary line (S1-S3, S1-S2 or S2-S3), the cluster can be propagated into the next cell. In this example if S3 detects CW rotation and S1 detects CCW rotation the cluster must be crossing into the A-S1-S3 region. Using this logic the cluster can be tracked throughout the sensor fabric as shown in **Figure 5.**

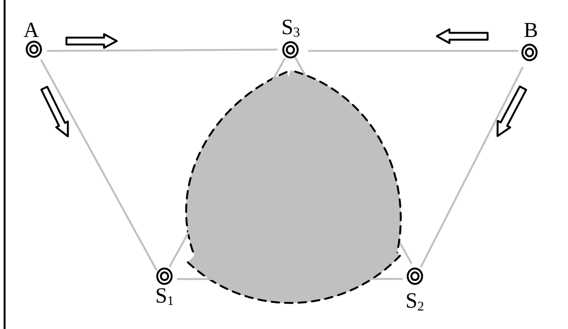

Figure 4: Acoustic arrays used as control volume sentries

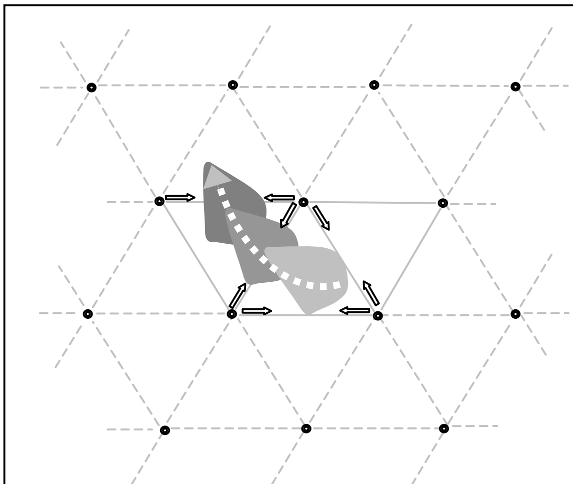

Figure 5: Movement of Cluster through a sensor fabric

Extension to higher levels

The distribution of potential targets (real and unresolved ghosts) within a region of acoustic activity will now be referred to as a "cluster". In general, the existence, size, shape, frequency content, and movement of For example, a cluster of lower-frequency clusters could provide useful information. sources that is growing in physical size, but otherwise appears stationary, could suggest the massing of tracked vehicles. Moreover, knowledge that a cluster has reorganized and is moving in the shape of an attack formation (Figure 5) could be just as important as the knowledge of individual vehicle locations within the cluster. Implementation of the above concept requires that the location of each sensor array is known, but it is not necessary that they be arranged in a regular grid pattern as depicted in Figure 5. In a more likely scenario, sensor arrays will be somewhat randomly distributed around the battlefield (Figure 6). Sensors that have become captured by local sources must be identified. Additionally, some provision should be made for continuity of cluster shape as a cluster moves from one vehicle-tracking sensor triad to another. Also, because sensor detection range is strongly affected by wind and terrain, the region of LOB intersections is not likely to be as symmetric as that shown for the isosceles triangle formed by S_1– S_2–S_3 in Figure 4. Thus some geometric compensation of detection boundaries and cluster shapes (perhaps using known wind data) may be required. Nevertheless, the principle of detecting boundary crossings should still be viable.

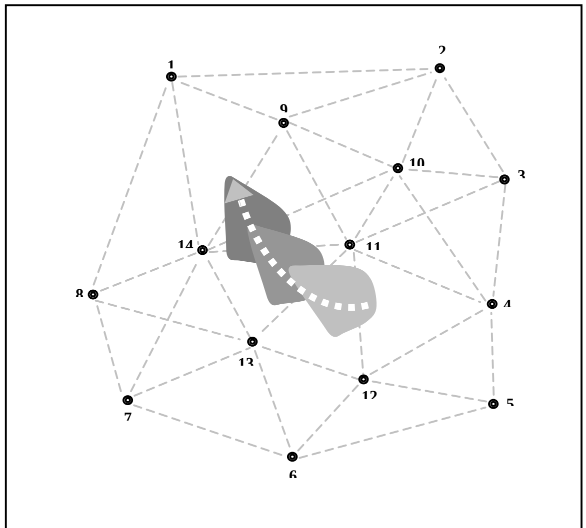

Figure 6: General case of distributed sensors

Cluster Shape Derivation. Returning to the scenario of **Figure 2**, let us now consider whether the arc-shape of the true target formation can be discerned from a 2-dimensional distribution of intersections. From the "cluster" definition stated above (i.e. targets and ghosts) the cluster shape when only two sensor arrays are functioning is just the exterior hull of all intersection points (**Figure 7).**

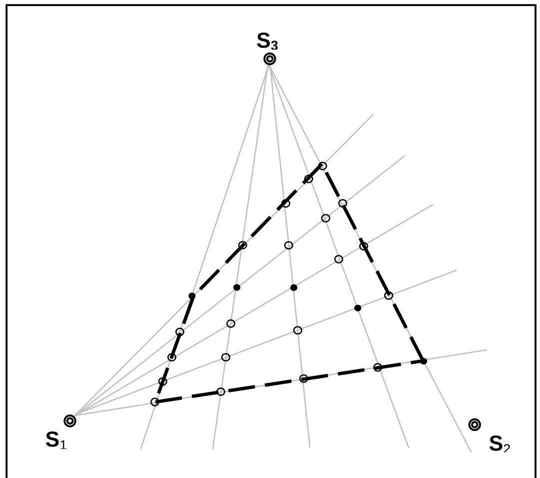

Figure 7: Cluster shape when Sensor S_2 is unavailable.

For the case where a third sensor is detecting all targets (**Figure 8a**), target T_1 is quickly found. (It's the only point that satisfies **S_3–5, S_2–5, and S_3–1**) While it *appears* that Target T_2 can be resolved as the average of the potential target intersection points along **S_3–2**, a similar averaging approach would fail completely along the line **S_2–4**. In that case, simple averaging would result in the predicted target location, **P_T**, illustrated in **Figure 8b**.

But what if the intersection distributions obtained from **S_3–2** and **S_2–4** are combined? Then, one would expect a "fuzzy" answer to be located somewhere between T_2 and P_T. Unfortunately, attempts to generalize this concept along each LOB (using convolution, area weighting, and Delaunay triangulation techniques) have been unsuccessful, resulting only in "fuzzy" convex blobs without useful interpretation.

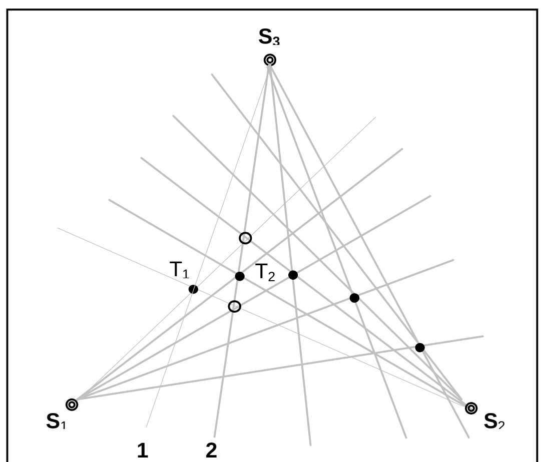

Figure 8a : T_2 appears as a simple weighted distribution along line S_3 – 2.

A better approach for deriving a cluster shape may be to test for the divergence of LOB intersections when potential targets *approach* a sensing region boundary (**Figure 9**). After errors associated with geometry are backed out, the area formed by the LOB intersections of true targets (**T_1 T_2 T_3**) should remain relatively constant between time intervals t=1 and t=2. But as ghosts (**G_1**, **G_2**) approach the boundary **$S_1 – S_3$** the associated areas should shrink or enlarge. Another possible way to "trim" ghosts from the cluster shape is to consider the kinematics of the potential targets as they *cross* a sensing region boundary.

From **Figure 10,** the true target **T_2** is *expected* to cross **$S_1 – S_3$** at the location where the radial and angular components of the total velocity vector **V** (i.e. V_R and V_T) are equal in magnitude relative to **S_1** and **S_3**. This implies that, at the boundary crossing**, $\mathbf{r_3 / r_1 = \omega_1 / \omega_3}$**. In general, since false targets will not meet this requirement, it should be possible to strip away ghosts by using LOB rate observations to estimate **r_3 / r_1** at the time of boundary crossing and then use that information in the target tracking filter. Importantly, by using forward-backward propagation of expected crossings, it may be possible to obtain a more reliable estimate of the vehicle count.

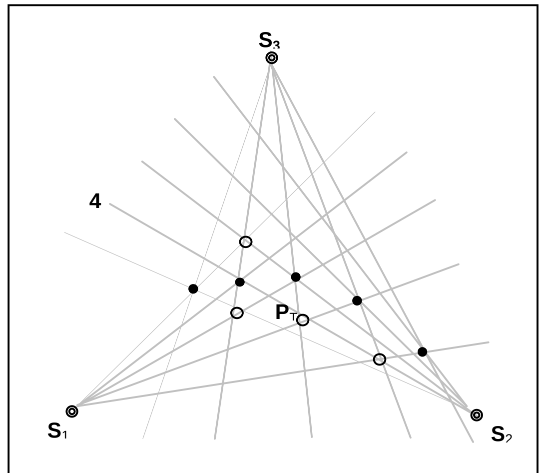

Figure 8b: Simple distribution of potential targets along line S_2–4 yields target P_T.

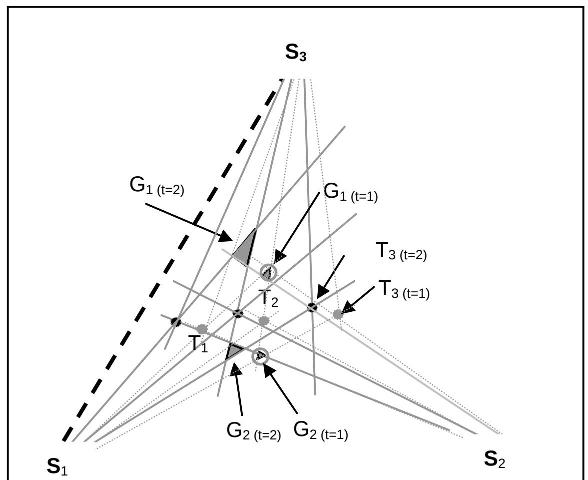

Figure 9: Movement of 3-LOB ghosts

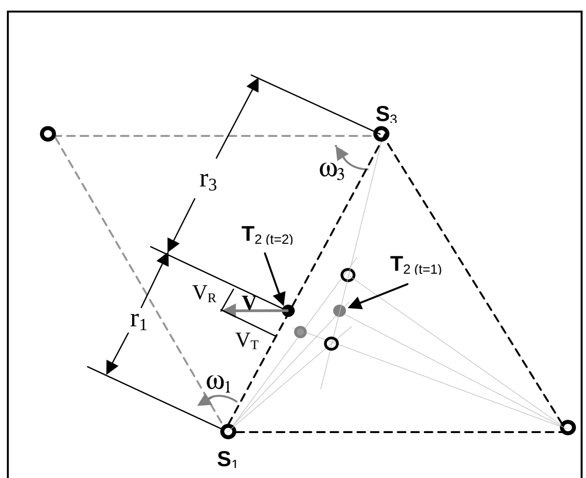

Figure 10: LOB dynamics used to trim ghosts from cluster shape

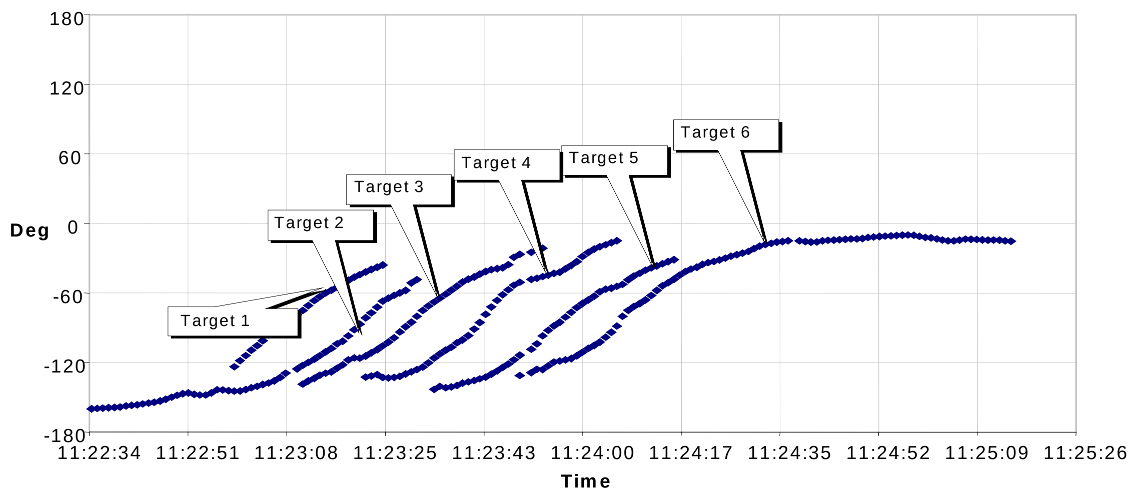

Figure 11. Real-time hardware tracking each of six vehicles in a closely spaced convoy (bearing angle vs. time). Vehicle separation is 50 meters. (From Reference 1)

Attention will now be given to an example of real acoustic sensor data provided by **Figure 11.** Several points are illustrated in this example: (1) Given sufficient spacing it is possible to resolve and count six targets. (2) While it is occasionally possible to see 4 simultaneous targets, usually only 3 targets can be discerned at any one time. (3) Inflections in the vehicle S-curves represent the closest distance between each vehicle and the sensor array. (4) Angular rate is provided by the slope of the S-curves. (5) Asymptotes exist at -10 degrees and -160 degrees. (6) A single sensor array may provide important information that should not be discarded just because vehicle locations cannot be computed (emphasizing the point made earlier).

For this specific example, Figure 11 states that the vehicle separation is 50 meters. Although this additional information may not be generally available, it can be used to construct a realistic geometry for the purpose of discussion. Assuming a straight column formation and a 50 meter separation distance, one can then deduce that the vehicle speed is 15.7 km/hr in a direction from -160 to -10 degrees. Also, the distance between asymptotes is about 490 meters and the closest distance between the vehicle path and the sensor array is about 65 meters. The questions that will now be asked are: (1) What will the S-curves look like to two other sensor arrays located to form a triad? (2) How can the individual S-curves be synchronized and propagated to boundary crossing points? And, (3) in the event that synchronization and propagation is not possible, can the information algorithm be gracefully degraded so the flux of vehicles between sensing regions continues to be monitored? Figure 12 has been constructed to answer these questions, where the real sensor $\mathbf{S_2}$ has been located at coordinate location (0,0) and sensors $\mathbf{S_1}$ and $\mathbf{S_3}$ have been added at coordinates (-200,-155) and (-200, 175) meters, respectively. **Figure 13** shows the calculated angles vs. time, as observed by $\mathbf{S_1}$ and $\mathbf{S_3}$, for the vehicle path in Figure 12. Notice that a boundary line drawn between $\mathbf{S_1}$ and $\mathbf{S_2}$ will make an angle of about 60 degrees relative to North. Similarly, a line between $\mathbf{S_3}$ and $\mathbf{S_2}$ will form an angle of 128 degrees. Both of these angles fall within the linear regions of the curves in **Figure 13** where changes in LOB-rates may not be observable. Consequently, if the LOB-rates obtained at the CPAs to $\mathbf{S_1}$ and $\mathbf{S_2}$

are used to predict the boundary crossing, then range errors Δr_{s1} and Δr_{s2} result as shown in **Figure 12.** Fortunately, for the case of constant linear velocity, the effect of these errors is negated by proportionality since:

$$\mathbf{r}_1 / \mathbf{r}_2 = (\mathbf{r}_1 + \Delta\mathbf{r}_{S1}) / (\mathbf{r}_2 + \Delta\mathbf{r}_{S2}) = \omega_2 / \omega_1$$

Moreover, it is *not necessary* for the LOBs to be recorded at the CPA. It is only necessary that the LOBs from $\mathbf{S_1}$ and $\mathbf{S_2}$ be parallel. This is important because it may not be possible to resolve which particular LOBs correspond to CPAs from single-sensor data.

Implementation. For the test data shown in **Figure 11** there are 6 targets with about 11 seconds of separation between them. As an example of how the S curves may be synchronized, assume that $\mathbf{S_1}$ first detects a LOB and a LOB-rate for an individual target whose range is unknown. A LOB-rate tracking filter is then initiated and is used to predict the time of boundary crossing along $\mathbf{S_1}$ -- $\mathbf{S_2}$. This prediction is continually updated as new data becomes available from $\mathbf{S_1}$. Once the predicted time of boundary crossing has occurred, a forward-time search is made of the data from a $\mathbf{S_2}$ LOB-rate tracking filter to "locate" (i.e. propagate within error limits) an LOB parallel to the most recent $\mathbf{S_1}$ LOB. If located, then the corresponding ratio of LOB-rates can be used immediately to compute $\mathbf{r_1}$ and $\mathbf{r_2}$ at the boundary crossing since,

$$\mathbf{r}_1 = (1 + \omega_{s1} / \omega_{s2})^{-1} \text{ (distance between sensors)}.$$

This information is also useful to the **target** tracking filter: If a potential **target** track (ghost or real) was *expected* at the boundary crossing point but is not confirmed by the $\mathbf{r_1}$ and $\mathbf{r_2}$ range determinations, then that track should be regarded as a ghost and terminated.

Conversely, the range measurement for a "confirmed target" can be used to update the **target** tracking filter with a data point that would otherwise be unobtainable through standard triangulation.

Although the preceding paragraph postulated that a LOB and LOB-rate were first observed by sensor $\mathbf{S_1}$, the order of detection is immaterial. If a LOB-rate track has been established at any sensor for a reasonable period of time (like the tracks shown in

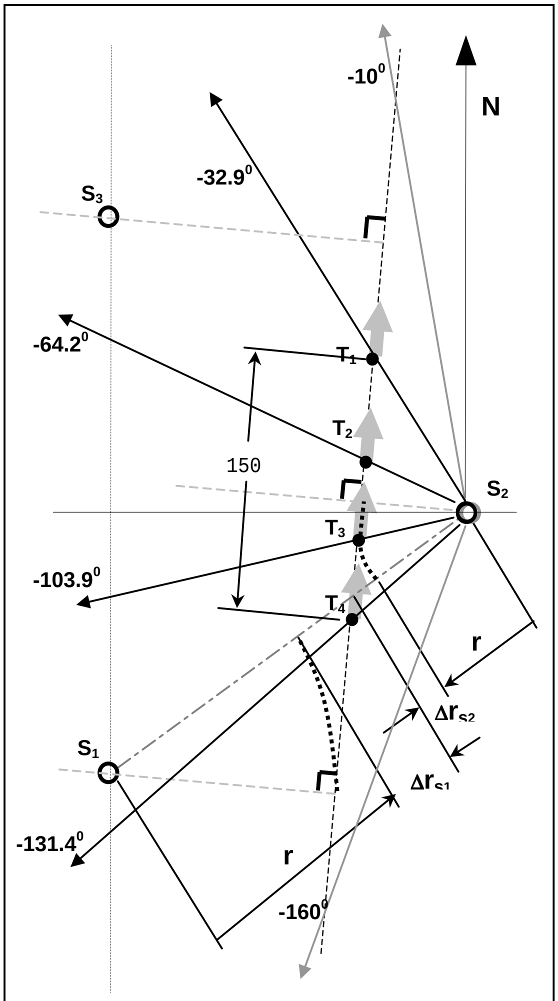

Figure 12: Target geometry at time = 11:23:25 , as reconstructed from figure 11.

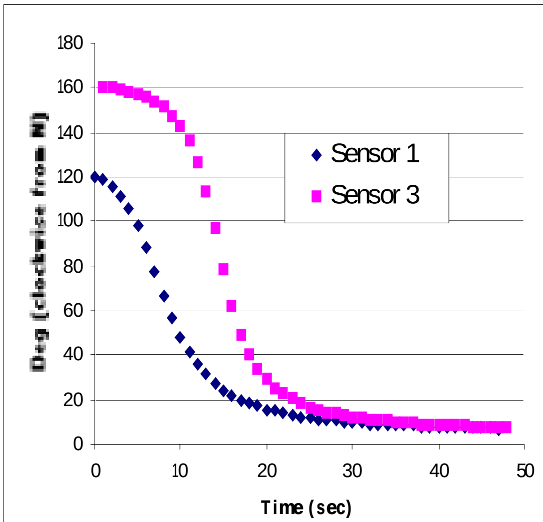

Figure 13: Target path observed by S_1 and S_3, (starting at callout 'A' in figure 12)

Figure 11) , then regardless of confirmation from another sensor's LOB-rate tracking filter, a vehicle counter should be incremented for the region of influx and decremented for the region of efflux at the time of the expected boundary crossing. To resolve potential counting issues, such as the double-booking of entries, it would be worthwhile to record both the LOB-rate tracker's sensor id and the times at which counter changes are made.

Discussion and Recommendations

This report has identified some technical problems which make it difficult for passive acoustic sensors to track individual targets on a battlefield. It goes on to suggest an important secondary role for such sensors.

A methodology has been presented which may make it possible to separate true and false targets as vehicles pass through a detection region, thereby culling the shapes of target formations and counting the number of vehicles contained therein. The methodology complements vehicle tracking algorithms in several respects. Positions established by kinematics at a detection boundary may be used to initialize new tracks or update existing tracks with data not available through triangulation. Moreover, false target elimination means that (1) greater measurement uncertainty can be tolerated in the vehicle tracking loops, and (2) multiple-hypotheses tracking algorithms using deeper history levels may become feasible. Better target tracking will in turn simplify the task of kinematic resolution; that it, by providing target positions closer to the detection boundary, a constant linear velocity approximation can be made more accurately and a search for parallel LOBs can be narrowed. It is recommended that algorithms be developed to test the proposed methodology using data acquired from prior field tests.

Conclusions

In applying acoustic sensor technology to the active battlefield, realistic sensor performance must be taken into consideration. Failure to due this will produce system level failures even though sensors are still producing valuable information. In the case of when the sensors are resolving multiple targets simultaneously the number of potential targets grows exponentially. Tracking all of the potential targets will quickly consume computational resources and ghost tracks can poses similar dynamics to real tracks precluding their rejection. This paper has presented concepts that allow the system level performance level to degrade gracefully even under such severe conditions of activity. The view of the problem at a micro level and a macro level ensures that the information produced by such a system has value. At the macro level information such as vehicle numbers, general directions, group amassment and group formation may be attainable. This theater level information may be just as valuable as the location of individual targets. Thus the primary goal is to allow the system to continue to provide such information when individual tracking at the micro level is not possible. Exploiting information about battlefield movements and battlefield target formations offers information not applicable to a general purpose tracking system. The use of such information may even change the primary measurables of a general acoustic sensor system to one that estimates the properties of a cluster of targets rather than the measurables to individual targets.

References

1. **WAM Acoustic Technical Feasibility Study**, N Srour, et al. Army Research Laboratory, 31 July 1988.